UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/057,331	01/24/2002	Andrei Z. Broder	5598/153US	1535

7590 02/08/2006

Seth Ostrow
BROWN RAYSMAN MILLSTEIN FELDER AND STEINER LLP
900 Third Avenue
New York, NY 10022-4728

EXAMINER	
HILLERY, NATHAN	

ART UNIT	PAPER NUMBER
2176	

DATE MAILED: 02/08/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/057,331	BRODER ET AL.	
	Examiner	Art Unit	
	Nathan Hillery	2176	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *22 November 2005*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-12 and 14-16* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-12 and 14-16* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

1. This action is responsive to communications: RCE filed on 11/22/05.

2. Claims 1 – 12, and 14– 16 are pending in the case. Claims 1, 9, and 16 are independent.

3. The rejection of claims 1 – 8 under 35 U.S.C. 101 as being non-statutory has been withdrawn.

4. The rejection of claims 9 – 12, and 14 – 16 under 35 U.S.C. 101 as being non-statutory has been maintained.

5. The rejection of claims 1 – 12, and 14 – 16 under 35 U.S.C. 103(a) as being unpatentable has been maintained.

Response to Amendment

The amendment to the claims filed on 10/27/05 does not comply with the requirements of 37 CFR 1.121(c) because Each amendment document that includes a change to an existing claim, cancellation of an existing claim or addition of a new claim, must include a complete listing of all claims ever presented. Amendments to the claims filed on or after July 30, 2003 must comply with 37 CFR 1.121(c) which states:

(c) *Claims*. Amendments to a claim must be made by rewriting the entire claim with all changes (*e.g.*, additions and deletions) as indicated in this subsection, except when the claim is being canceled. Each amendment document that includes a change to an existing claim, cancellation of an existing claim or addition of a new claim, must include a complete listing of all claims ever presented, including the text of all pending and withdrawn claims, in the application. The claim listing, including the text of the claims, in the amendment document will serve to replace all prior versions of the claims, in the application. In the claim listing, the status of every claim must be indicated after its claim number by using one of the following identifiers in a parenthetical expression: (Original), (Currently amended), (Canceled), (Withdrawn), (Previously presented), (New), and (Not entered).

Continued Examination Under 37 CFR 1.114

6. A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on 10/27/05 has been entered.

Claim Objections

7. Claim 15 is objected to because of the following informalities: depends from a cancelled claim. Appropriate correction is required.

Claim Rejections - 35 USC § 103

8. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

9. Claims 1 – 12 and 14– 16 are rejected under 35 U.S.C. 103(a) as being unpatentable over Bharat et al. (US 6112203 A).

10. **Regarding independent claim 1**, Bharat et al. teach that *the set of documents can be produced by combining the set of results from a Web search engine in response to a user query (which we call the `start-set`), with pages that either link to or are linked from the start-set documents ... The nodes in the start set are first scored according to their connectivity, and the number of terms of the query that appear as unique sub-*

strings in the URL of the represented documents. The score is a weighted sum of the number of directed edges to and from a node and the number of unique sub-strings of the URL that match a query term (Column 3, lines 3 – 15), which provide for **receiving a document to be processed; locating a set of documents that include hyperlinks to the document; retrieving anchortext** (*sub-strings in the URL*) **associated with at least one of the hyperlinks**, and **parsing the anchortext** (*sub-strings in the URL*) **into one or more tokens**. Bharat et al. also teach that *as previously disclosed, uses the relevance weights of all of the nodes to decide whether or not to eliminate a page for user consideration. For example, prune all nodes whose relevance weight is below a predetermined threshold. The threshold can be picked in a number of ways* (Column 6, lines 22 – 27), which provide that **for each token: determining a weight for the token, determining whether the weight assigned to the token exceeds a threshold token weight**. Bharat et al. do not explicitly teach **indexing**. However, it would have been obvious to one of ordinary skill in the art to be motivated to use and/or alter the invention of Bharat et al. to provide for **indexing the document under the token, if the token weight assigned to the token exceeds the threshold token weight**, since Bharat et al. do teach that *in order to help users locate Web pages of interest, a search engine 140 maintains an index 141 of Web pages in a memory, for example, disk storage* (Column 4, lines 9 – 11) and that *we provide an improved ranking method 200 that can be implemented as part of the search engine 140. Alternatively, the method 200 can be implemented by one of the clients 110, or some other computer system on the path between the search engine and the clients* (Column 4, lines 23 – 27).

11. **Regarding dependent claims 2, 5, 10 and 11**, Bharat et al. teach that *specifically, in step 220, we score each page p of the input set 201 to determine a value Score(p) 225. Let n_p be the node representing page p. The score is determined by: Score(p)= in_degree + 2 x (num_query_matches) + out_degree, where in_degree is the number of edges pointing at node n_p, num_query_matches is the number of unique sub-strings of the URL of the page p that exactly match a term in the user's query* (Column 5, lines 57 – 64), which provide for **including in the index an indication of weight for each token under which each page is indexed**, and that **the weight of each token is based on its frequency of occurrence within the index**.

12. **Regarding dependent claims 3 and 4**, Bharat et al. teach that *next, we assign a relevance weight to a subset of the nodes 212. The relevance weight measures the similarity between the represented page and the query topic. As stated above, the topic implied by the user is probably broader than the query itself. Thus, matching the words of the query with the page is usually not sufficient. Instead, as described in detail below, we use a subset of the pages of the start set 201 to define a broader query topic "Q", and match the pages "P" represented in the graph with the broader query topic to determine the relevance weights of the nodes 212. Our invention is motivated by the observation that not all pages represented by nodes in the n-graph 211 are equally influential in deciding the outcome of our ranking process* (Column 5, lines 21 – 33). Bharat et al. do not explicitly teach assigning the token to a particular location; however, one of ordinary skill in the art at the time of the invention would be motivated to alter the invention of Bharat et al. to provide for **assigning to the token a location within the**

index corresponding to part of the page being indexed that is allocated for tokens having a higher degree of importance than other tokens in the same page, and for **assigning to the token a location within the index that corresponds to the beginning of the page being indexed**, since the skilled artisan would want to point the user to the exact location in the page or to the beginning of the page so that the user does not have to hunt for the exact location himself or become confused and/or overwhelmed by the results and information that he is trying to interpret.

13. **Regarding dependent claims 6 and 7**, Bharat et al. teach that *because the query topic Q 245 can include a large number of terms, and because the "vocabulary" of the various pages can vary considerably, we prefer to use term frequency weighting. More specifically, we use cosine normalization in weighting both the query topic Q and the pages P because the deviation in term vector lengths is large, specifically: ... where* $w_{iq} = freq_{iq} \times IDF_i$, $w_{ij} = freq_{ij} \times IDF_i$, $freq_{iq}$ *is the frequency of (stemmed) term i in the query topic Q,* $freq_{ij}$ *is the frequency of term i in page j, and* IDF_i *is an estimate of the inverse document frequency (IDF) of the term i in the corpus of documents, for example, in our case, a large representative sample of Web pages* (Column 7, lines 10 – 29), which provide for **determining a first frequency at which the anchortext appears in the index; determining a second frequency at which each token derived from the anchortext appears in the index; and assigning a weight to the token, wherein the weight is a function of the first and second frequencies**, and **dividing the first frequency by the second frequency to produce a weight**

quotient; and multiplying the weight quotient by an anchor text count for the token.

14. **Regarding dependent claim 8**, Bharat et al. teach that *during a connectivity analysis phase, the remaining nodes of the pruned graph are then scored according to their connectivity to determine normalized hub and authority scores for the documents. The normalized scores are used to rank the documents* (Column 3, lines 31 – 35), which provide for **determining a normalized weight for each token**.

15. **Regarding dependent claims 12 – 14,** the claims incorporate substantially similar subject matter as claims 4 – 6, and is rejected along the same rationale.

16. **Regarding dependent claim 15**, the claim incorporates substantially similar subject matter as claim 8, and is rejected along the same rationale.

17. **Regarding independent claims 9 and 16**, the claim incorporates substantially similar subject matter as claim 1, and is rejected along the same rationale.

Response to Arguments

18. Applicant's arguments filed 10/27/05 have been fully considered but they are not persuasive.

19. In response to Applicant's arguments that *anchortext is the text associated with a URL, not the URL itself* (p 7, lines 16 – 17), it should be noted that the passage within the specification on p 2, lines 11 – 18 as cited by Applicant does not exclude anchortext from being the text of the actual URL; however, the passage does state that *In the page containing the link, usually there is some text associated with the link. In typical browsers the user clicks on this text to follow the link. This text is known as anchortext.*

In other words, anchortext is the text a user clicks to follow a link. Within the broadest, reasonable interpretation, the term anchortext is a term of art known to those of ordinary skill as evidenced by the Google definitions cited by the Office. Google shows that while anchortext can be non-URL text such as "linking text or anchor text" (p 1, under Web), it can also be the text of the URL such as "www.patrickgavin.com/SEO-Glossary.htm" (p 1, under first definition) in accordance with the first definition *Also known as Link Text, the clickable text of a hyperlink* (p 1). Furthermore Google defines linking text or anchor text as simply *The text that is contained within a link* (p3). Further by applicant's own admission the anchortext associated with the URL "www.UAL.com" may be any combination of characters (p 8, lines 3 – 4), which can include the text of the URL and/or any substrings thereof as evidenced by *The anchortext associated with the link might be "United", "United Air Lines," or* ***"U.A.L,"*** *entirely at the discretion of the author of the page linking to the United site* (Specification, p 2, lines 11 – 18) and disclosed by Bharat et al. (Column 3, lines 3 – 15) as outlined above in the rejection of the claims.

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Nathan Hillery whose telephone number is (571) 272-4091. The examiner can normally be reached on M - F, 10:30 a.m. - 7:00 p.m.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Heather R. Herndon can be reached on (571) 272-4136. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

NH

HEATHER R. HERNDON
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100